REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-01-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
6–8–2001	Technical	

4. TITLE AND SUBTITLE

INFORMATION ASSURANCE CONSIDERATIONS FOR A FULLY NETTED FORCE: IMPLEMENTING CRANOF FOR STRATEGIC INTRUSION ASSESSMENT FOR CYBER COMMAND AND CONTROL

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
0601152N

6. AUTHORS

I. R. Goodman W. C. Torrez
D. Bamber Naval War College
SSC San Diego

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

SSC San Diego
San Diego, CA 92152-5001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217–5000

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

We consider here a first step in applying CRONOF (Complexity Reducing Algorithm for Near Optimal Fusion—a new rigorously based complexity-reducing algorithm that produces estimates of undereonstrained probabilities—to the problem of determining when a computer network is under attack. Essentially, CRANOF treats this issue as a formal analogue of the pairwise track correlation or similarity problem, comparing the current cyber-state history with each of various alternative classes of cyber-state histories relative to various features or attributes measuring various degrees of normality/abnormality.

Published in *Proceedings of the Fourth International Conference on Information Fusion* (FUSION 2001), vol. 2, 2002, ThC3-19–ThC3-24.

15. SUBJECT TERMS

Mission Area: Information Science
information assurance intrusion transitivity
cyber-states second-order probability

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			I. R. Goodman
U	U	U	UU	7	19B. TELEPHONE NUMBER *(Include area code)* (619) 553-3680

INFORMATION ASSURANCE CONSIDERATIONS FOR A FULLY NETTED FORCE:

IMPLEMENTING CRANOF FOR STRATEGIC INTRUSION ASSESSMENT FOR CYBER COMMAND AND CONTROL

Dr. William C. Torrez

CNO STRATEGIC STUDIES GROUP

Newport, Rhode Island 02841
mailto:torrezw@nwc.navy.mil

Dr. Donald Bamber and Dr. I. R. Goodman

SPACE AND NAVAL WARFARE SYSTEMS CENTER,
Code D44215

San Diego, California 92152
mailto:bamber@spawar.navy.mil
mailto:goodman@spawar.navy.mil

Paper published in Proceedings of the Fourth International Conference on Information Fusion (FUSION 2001), held August 6-10, 2001, Montréal, Québec, Canada: volume 2, pages ThC3-19 – ThC3-24.

INFORMATION ASSURANCE CONSIDERATIONS FOR A FULLY NETTED FORCE:
IMPLEMENTING CRANOF FOR STRATEGIC INTRUSION ASSESSMENT AND CYBER COMMAND AND CONTROL

William C. Torrez
CNO STRATEGIC STUDIES GROUP,
NAVAL WAR COLLEGE
Newport, Rhode Island 02841
torrezw@nwc.navy.mil

Donald Bamber and I. R. Goodman
SPACE AND NAVAL WARFARE SYSTEMS
CENTER, Code D44215
San Diego, California 92152
bamber@spawar.navy.mil
goodman@spawar.navy.mil

Abstract – *We consider here a first step in applying CRANOF (Complexity Reducing Algorithm for Near Optimal Fusion – a new rigorously-based complexity-reducing algorithm that produces estimates of underconstrained probabilities – to the problem of determining when a computer network is under attack. Essentially, CRANOF treats this issue as a formal analogue of the pairwise track correlation or similarity problem, comparing the current cyber-state history with each of various alternative classes of cyber-state histories relative to various features or attributes measuring various degrees of normality / abnormality.*

Keywords - Information assurance, cyber-states, intrusion, second order probability, transitivity.

1. Introduction

For a fully netted force in the years 2010 and beyond, information systems will constitute a critical "center of gravity" and must be designed to be survivable. Fortunately, future Network-Centric Warfare (NCW) concepts will depend on widely dispersed network nodes that make a "hard-to-find" center of gravity. This naturally survivable and gracefully degradable architecture will still need an active and effective resident Information Assurance (IA) capability. NCW networks and related systems must be robust and able to absorb faults and intrusions without significant reductions in capability. While it cannot be assumed that IA will make NCW unassailable in 2010, IA will insure that NCW systems are able to deliver the capabilities required by Naval Power Forward [18]. Activities such as Information Assurance, computer network defense and counter-deception will defend decision-making processes by neutralizing an adversary's perception management and intelligence collection. Two important technologies in Information Assurance are Strategic Intrusion Assessment (SIA) and Cyber Command and Control (CC2).

The Common Intrusion Detection Framework (CIDF) working group has stated [8] that two key problems in SIA are (a) the fusing and correlating of event and sensor information and (b) the tracking of attacks. Furthermore, in the CC2 program, a key problem is developing situation awareness. These three problems are all, of course, problems in higher levels of fusion. It is also our thesis that a close analogy exists between the problem of track-to-track correlation of kinematic targets and the problem of situation awareness via fusion of attack information about computer networks. In this analogy, the concept of alternative track histories corresponds to the concept of alternative cyber-state histories. By "cyber-state history", we mean either a description of past attacks of various kinds, or the temporal patterns expected to be observed in various types of unprecedented attacks, or even non-attack disruptions. Consider then the general problem of cyber-attack classification and fusion comparing the current cyber-state history with each of various alternative classes of cyber-state histories. In framing this problem, one can employ three basic types of random vectors, with corresponding conditional probabilities representing associated uncertainties: (a) correlation/similarity levels, (b) cyber-attribute (or cyber-feature) matching (or non-matching) of attributes and (c) observed cyber-attribute data from each cyber-state history. The correlation level variable represents the degree of similarity between the current cyber-state history and alternative cyber-state histories representing various attack and non-attack disruptions. The goal is to estimate the conditional probability of various correlation levels, given the observed data, under circumstances where the joint probability measure of the basic random variables is only partially obtainable.

2. Underconstrained Probability Problem

2.1 Cyber intrusion detection as a special case of the matching problem

Strategic intrusion assessment for cyber command and control obviously depends upon judgment (human, mechanistic, or some combination of both) that the degree of similarity between the current cyber state of affairs and any of a prescribed set of states during past intrusions – or even perceived future intrusions – is sufficiently high as to warrant corresponding defensive action. As stated in the Introduction, the above problem is analogous to that of track correlation, where a number of geolocation and non-geolocation attribute estimates in error are considered for matching in order to determine whether or not the two tracks represent the same target of interest or not. (For an example of an approach prior to the use of CRANOF, see, e.g., [19].)

Analogies can be established with other various pattern matching problems, including fingerprint identification, photography matching, use of clues left at crime scenes, and everyday recall of situations sufficiently similar to past events. Furthermore, this large class of matching problems is yet a special case of an even larger class of problems: the underconstrained probability class. To see this, consider the following problem expresses in terms of conditional statements that need only be partially true.

Given:
"If attribute j for X and Y matches, for j = 1,..., m, then X and Y are the same";
"If observe (in possible error) X_j for X and Y_j for Y (wrt attribute j), then X and Y match", j = 1,..., m;
Determine:
"If observe X_j for X and Y_j for Y, j = 1,..., m, then X and Y are the same" (1)

In terms of corresponding conditional probability evaluations and assuming, for simplicity, mutual independence of each of the conditionals in the second set of expressions:

$$P(a \mid b) = s \ , \quad P(b \mid c) = t, \tag{2}$$

where, using standard boolean algebra terminology,

a = "X and Y are same type of cyber attack / disruption",
$b = b_1 \& \ldots \& b_m$,
b_j = "X and Y match wrt attribute j", j = 1,..., m,
$c = c_1 \& \ldots \& c_m$, $c_j = (X_j, Y_j)$, j = 1,..., m, (3)

and

$$P(b|c) = P(b_1|c_1) \cdots P(b_m|c_m). \tag{4}$$

Thus, in summary, the matching problem – in quite simplified form – can be phrased as

$$\text{Given } P(a|b) = s, P(b|c) = t; \text{ Determine } P(a|c), \tag{5}$$

where s and t are known either exactly or approximately and where a, b_j, b, c_j, c are all as in eq.(3).

While commonsense reasoning indicates that in general, if s and t are reasonably high (i.e., close to unity), then so should P(a|c), *it does not actually follow that a particular P will satisfy this property.* This *transitivity* problem, which is a probabilistic version of the famous Aristotle syllogism – where, in terms of probability formulation, if s = t = 1, the conclusion P(a|c) = 1 indeed holds – has provoked great controversy in the AI community. In fact, one can readily construct probability measures P such that s, t can be arbitrarily close to unity (but not exactly unity), with P(a|c) close to, or even equal to zero. (For background on this issue, see, e.g., Pearl [17].) Thus, "determine P(a|c)" in eq.(4) should be replaced by "estimate P(a|c) in some best sense":

Given P(a|b) = s, P(b|c) = t;
Estimate P(a|c) in some best sense, (6)

If the above independence assumption leading to eq.(4) is felt to be unwarranted, but at least the marginal conditional probabilities $P(b_j|c_j)$ are known or estimable, then the matching problem can be modified – and relaxed, appropriately – so that eq.(6) is replaced by

Given P(a|b) = s, $P(b_j|c_j) = t_j$, j = 1,..., m;
Estimate P(a|c) in some best sense, (7)

assuming s, t_j are obtainable, j = 1,..., m.

More generally, the underconstrained probability problem can be phrased as

Given: $P(a_1|b_1) = t_1, \ldots, P(a_m|b_m) = t_m$;
Estimate in some best sense P(e|f). (8)

Eq.(8) also arises in problems involving rule-based systems – where the given (not necessarily perfect) rules are of the form "if b_j, then a_j", the given (possibly partially true) facts are $f_1, \ldots, f_n$, and it is desired to ascertain whether e is also a fact. In this case, one simply interprets f as the conjunction $f_1 \& \ldots \& f_n$.

Many other patterned special cases can also be considered either as extensions of classical logic valid or invalid entailment schemes or arising from AI considerations. (See, e.g., [14] for more details.) On the other hand, underconstrained probability problems need not initially be connected with any logical entailment considerations, but simply arise from probabilistic models describing a multitude of military problems, including surveillance, search, detection, prediction, and reliability, among many others. All that is required is that the conditional probability of interest P(e|f) not be uniquely – or inconsistently – specified by the known probabilities of the given or premise collection of conditional probabilities. Also, note that a basic alternative form of the *exact threshold formulation* of the estimation problem presented in eq.(8) is the corresponding *lower bound threshold formulation*:

Given: $P(a_1|b_1) \geq t_1, \ldots, P(a_m|b_m) \geq t_m$;
Estimate in some best sense P(e|f). (9)

Again, as in the transitivity and modified transitivity problem, the events a_j, b_j, e, f may all be related in some patterned way, or perhaps in no particular way.

Finally, a basic related issue to that in eq.(8) is to estimate P(e|f) in some sense when the thresholds are made to approach unity in some sense (such as uniformly or at various prescribed rates)

2.2. Estimation aspect of CRANOF

One approach to estimating the desired conclusion probability P(e|f) in eq.(9) – as originated with Adams (and usually provided in more general conditional form to be discussed below) [1, 2] – is to take a pessimistic viewpoint in selecting some set of fixed lower bound probability thresholds t_j corresponding to each premise event a_j, j = 1,..., n, and determine as a function of $\underline{t} = (t_1,..., t_n)$, the *minimum conclusion function to at least degree* $\underline{t}$,

$$\text{minconc}(\{\text{if } b_1, \text{ then } a_1,..., \text{ if } b_m, \text{ then } a_m\}; \text{ if f, then e})(\underline{t})$$
$$= \inf \{P(e|f): P \text{ is a probability measure over } B \text{ and } P(a_j|b_j) \geq t_j, j = 1,..., m\}. \quad (10)$$

While, at first glance, the use of the minimum conclusion function appears reasonable, it can be shown that), for all thresholds $\underline{t}$ sufficiently close to, but distinct from, unity: (a) A number of key reasoning schemes fitting the general format of eq.(9) with "best estimate" interpreted via eq.(10) lead to the trivial value of 0. This includes transitivity [16, 3, 4]. (b) For *any* reasoning scheme, the limiting value of the estimate of P(e|f) is either 0 or 1, thus not allowing for nontrivial "degrees of validity or confidence". (Again, see [1, 2, 4, 5] for details.)

A less pessimistic approach – that of CRANOF –addresses the general underconstrained probability problem as explicated in Section 2.1, taking into account both optimality of solution and implementation complexity. Here, one replaces the minimum conclusion function above by the *mean conclusion function* [4, 5], where for some choice of prior distribution D over possible P's of interest (such as corresponding to a uniform distribution or various biased distributions), indicating, as usual, conditional expectation with respect to choice D of P's as $E_D(.|..)$,
the lower bound threshold formulation is

$$\text{meanconc}_1(\{\text{if } b_1, \text{ then } a_1,..., \text{ if } b_m, \text{ then } a_m\}; \text{ if f, then e})(\underline{t})$$
$$= E_D(P(e|f) \mid P \text{ is a probability measure over } B \text{ and } P(a_j|b_j) \geq t_j, j = 1,..., m), \quad (11a)$$

with the obvious analogue holding for the exact threshold formulation

$$\text{meanconc}_2(\{\text{if } b_1, \text{ then } a_1,..., \text{ if } b_m, \text{ then } a_m\}; \text{ if f, then e})(\underline{t})$$
$$= E_D(P(e|f) \mid P \text{ is a probability measure over } B \text{ and } P(a_j|b_j) = t_j, j = 1,..., m). \quad (11b)$$

The conditional expectation(s) in eq.(11) must be further explained to make sense. First, consider the boolean algebra B of all events naturally generated from the key components making up the conditional expressions "if b_j, then a_j", j = 1,..., m, as well as "if f, then e", or equivalently, the conditional probabilities $P(a_j|b_j)$, j =1,..., m, and P(e|f): i.e., from the class of events

$$C = \{a_j \& b_j, b_j: j= 1,..., m\} \cup \{e\&f, f\}. \quad (12)$$

Define, for ease of notation here,

$$a_0 = e, b_0 = f, \quad (13)$$

and for each j, define

$$\omega_{j,1} = a_j \& b_j; \ \omega_{j,2} = a_j' \& b_j, \ \omega_{j,3} = b_j'. \quad (14)$$

In turn, consider the class G of functions g, where

$$G = \{g: g \text{ maps } \{0,1,...,m\} \text{ into } \{1, 2, 3\}\} = \{1, 2,3\}^{\{0,1,...,m\}} \quad (15)$$

and for each g in G, define the *relative atom* ω_g *determined by* g *acting on* C as

$$\omega_g = \omega_{0,g(0)} \& \omega_{1,g(1)} \& ... \& \omega_{m,g(m)}. \quad (16)$$

Next, define the class of all such nonvacuous relative atoms as

$$A = \{\omega_g: g \text{ in } G \text{ and } \omega_g \neq \varnothing\}. \quad (17)$$

Then, it is easily verified that B consists of all finite disjoint disjunctions (v) of ω_g in A. Hence, any probability measure P that is well defined with respect to the conditional expressions making up the problem in eqs.(8) or (9) – or equivalently, well-defined upon the key events making up C – is uniquely determined by its values over A. In fact, P can be naturally identified as a probability function over A, or equivalently, as the q by 1 probability vector (indicating matrix or vector transpose by $(.)^T$)

$$\underline{P}^T = (P(\omega_g))_{g \text{ in } A}; \ q = \text{cardinality}(A). \quad (18)$$

In turn, totally ordering all q relative atoms and identifying them as element 1,..., element q, any prior second order probability distribution D of probability measures P for the problem addressed in eqs.(8) or (9) can be naturally identified as an ordinary probability distribution over the space of possible $\underline{P}$'s in eq.(18), i.e., over the q-simplex in surface form

$$S_q = \{\underline{P}: \underline{P}^T = (p_1,....p_q): 0 \leq p_j \leq 1, p_1 +...+ p_q = 1\}, \quad (19)$$

or equivalently, over the full (q-1)-dimensional simplex

$$S_{(q)} = (\underline{P}: \underline{P}^T = (p_1,....p_{q-1}): 0 \leq p_j \leq 1, p_1 +...+ p_{q-1} \leq 1\}, \quad (20)$$

One natural choice of prior D over $S_{(q)}$ is the uniform one. More generally, it can be shown that in a natural sense, the optimal choice of family of possible priors D is that of the Dirichlet form [12]. In any case, the above argument demonstrates explicitly how the conditional expectation in eq.(9) makes sense. It also follows that the optimal estimate of P(e|f), which is obviously the expectation of the bayesian posterior, in general will be uniquely achieved, and from classical results, due originally to Wald [20], will therefore be a decision-theoretic admissible estimator relative to the usual norm-

square loss function. The robustness properties of such posterior estimators are also well-known [13, 16].

Another approach to the interpretation of "estimate in some best sense" in eqs.(8) or (9) is the use of maximal entropy [9]. Alternatively, the term "estimate in some best sense" in these equations can be avoided altogether by taking an upper/lower probability bound approach [7, 21]. (Ongoing research is being conducted by the authors investigating the relationship between the bayesian posterior approach of the mean conclusion function with these other approaches.)

Returning to eqs.(8), (9), one can first show the basic relation between the exact and lower bound formulations is that the latter is a weighted integral of the former and that the former can be expressed succinctly as the ratio of two well-defined surface integrals, that, in turn, can be evaluated as ordinary integrals [10], improving and extending earlier results.

Often, it is more convenient to consider in place of either meanconc$_j$ function in eq.(11), the natural corresponding *plug-in* forms

$$\text{meanconc}_3(\{\text{if } b_1, \text{ then } a_1, \ldots, \text{ if } b_m, \text{ then } a_m\}; \text{ if f, then e})(\underline{t}) = P^{\#}{}_3(e|f) ; \quad (21a)$$

$$P^{\#}{}_3 = E_D(P \mid P \text{ is a probability vector in } S_{(q)} \text{ and } P(a_j|b_j) \geq t_j,\ j = 1,\ldots, m); \quad (22a)$$

$$\text{meanconc}_4(\{\text{if } b_1, \text{ then } a_1, \ldots, \text{ if } b_m, \text{ then } a_m\}; \text{ if f, then e})(\underline{t}) = P^{\#}{}_4(e|f) ; \quad (21b)$$

$$P^{\#}{}_4 = E_D(P \mid P \text{ is a probability vector in } S_{(q)} \text{ and } P(a_j|b_j) = t_j,\ j = 1,\ldots, m), \quad (22b)$$

noting that by a straightforward convexity argument, both $P^{\#}{}_3$ and $P^{\#}{}_4$ are also probability vectors in $S_{(q)}$.

2.3. Asymptotic and complexity reduction aspect of CRANOF

The actual evaluation of the seemingly simple-appearing integrals mentioned at the end of subsection 2.2 can, in general, be very computationally intensive, due to the possible presence of a large number (m) of premise set constraints $P(a_j|b_j) = t_j$. On the other hand, if it is reasonable to assume that all of the thresholds are sufficiently close to unity, relatively simple results have been obtained by Bamber under the assumption that prior D corresponds to a uniform distribution over S_q [3]. Returning to the non-limiting threshold case, a basic justification has been derived for an approximation to meanconc whereby the original premise class is replaced by a single (albeit complex) rule that for the asymptotic threshold case provides equivalent behavior of meanconc. This requires a certain sufficiency condition to be satisfied ([4], section 6). (For a basic formulation and application to the modified transitivity problem, see Section 3 below.)

3. CRANOF applied to transitivity and modified transitivity

3.1. Basic Results

First, it should be remarked that a related transitivity-like application of CRANOF to track correlation is provided in [6].

Returning to the more specific class of cyber-state problems considered here, that of transitivity in eq.(6) can be fully evaluated by specializing the previously discussed surface integral techniques [10], assuming D corresponds to a uniform distribution over $S_{(q)}$. Here, q = 7, since it is readily shown here that the conditionals "if c, then b", "if b, then a", and "if c, then a" lead to A = {a&b&c, a&b&c′, a&b′&c, a′&b&c, a′&b&c′, a′&b′&c, b′&c′} and hence card(A) = 7:

$$\text{meanconc}(\{\text{if c, then b, if b, then a; if c, then a}\})(s,t) = E_D(P(a|c): P \text{ is a probability measure over } B \text{ and } P(a|b) = s,\ P(b|c) = t) = s{\cdot}t + (1-t)/2 - h(s,t) ; \quad (23)$$

where

$$h(s,t) = h_1(s,t) / h_2(s,t);\ h_1(s,t) = s{\cdot}(2s-1){\cdot}(1-s){\cdot}t{\cdot}(1-t^2);$$
$$h_2(s,t) = t + 2t^2 + h_3(s,t);\ h_3(s,t) = s{\cdot}(1-s){\cdot}(1-t){\cdot}(2+3t-t^2). \quad (24)$$

For s, t reasonably close to unity, one can use the simplifying approximation

$$h(s,t) \approx (2/3){\cdot}(1-s){\cdot}(1-t). \quad (25)$$

The solution to the above problem can be modified, where probabilistic inputs are replaced by linguistic population-conditioned ones. (For details, see [11].)

While closed-form results have been obtained for not only transitivity above in eqs.(23), (24), but a whole host of other types of estimation schemes stemming from classical logic and rule-based system considerations (such as *contraposition, positive conjunction, cautious monotonicity, abduction, strengthening of antecedent*, etc. – again, see again [4]), many other more complicated arguments cannot be so readily obtained. But, as mentioned earlier, if a certain sufficiency condition is satisfied, then the argument in question can, in an asymptotic equivalent threshold sense, be replaced by a singleton premise class, which, in turn, leads to a closed-form evaluation.

Theorem 1. (Bamber & Goodman [4].)
Consider the problem in eq.(8), addressed by use of the meanconc function in eq.(11b). Suppose that the prior distribution D has a probability density function over $S_{(q)}$ that is continuous and uniformly bounded away from both zero and infinity and that the condition

$$a_j\&(b_1 \Rightarrow a_1)\&\ldots\&(b_m \Rightarrow a_m) \neq \varnothing, \text{ for } j = 1,\ldots, m \quad (26)$$

holds, where the material implication operator $\Rightarrow$ is defined, as usual, to be of the form

$$b_i \Rightarrow a_i = b_i' \vee a_i = b_i' \vee a_i\&b_i . \quad (27)$$

Then, letting here $\underline{t} = (t, t, \ldots, t)$,

(i)

$$\lim_{t \uparrow 1} [\text{meanconc}_3(\{\text{if } b_1, \text{ then } a_1, \ldots, \text{ if } b_m, \text{ then } a_m\}; \text{ if f, then e})(\underline{t})]$$

$$= \lim_{t \uparrow 1} [\text{meanconc}_4(\{\text{if } b_1, \text{ then } a_1, \ldots, \text{ if } b_m, \text{ then } a_m\}; \text{ if f, then e})(\underline{t})]$$

$$= \lim_{t \uparrow 1} [\text{meanconc}_2(\{\text{if } \beta, \text{ then } \alpha\}; \text{ if f, then e})(t)] ; \quad (28)$$

$$\beta = b_1 \vee \ldots \vee b_m \, ; \, \alpha = \beta \& (b_1 \Rightarrow a_1) \& \ldots \& (b_m \Rightarrow a_m). \quad (29)$$

(ii) Analogous relations hold for the limiting forms of meanconc_j for $j = 1, 2$ ■

Theorem 2. (Bamber & Goodman [4, 6])
Make the same assumptions as in Theorem 1, now with D chosen as $D(\underline{\tau})$, where $D(\underline{\tau})$ is the Dirichlet probability distribution with parameter vector $\underline{\tau} = (\tau_1, \ldots, \tau_q) \geq (1, 1, \ldots, 1)$, and with α and β being as in eq.(29). Suppose, without loss of generality, that the class of relative atoms A, defined in (17), has been indexed so that,

$$A = \{\omega_1, \ldots, \omega_{q-1}, \omega_q\}, \quad (30)$$

$$\omega_q = \beta' \& f' = b_1' \& \ldots \& b_m' \& f'. \quad (31)$$

For any nonvacuous event c in B, the boolean algebra determined by A, let I(c) denote the unique minimal index set, $\varnothing \neq I(c) \subseteq \{1, \ldots, q\text{-}1, q\}$ such that

$$c = \vee_{j \text{ in } I(c)}(\omega_j) \text{ (disjoint disjunction).} \quad (32)$$

Correspondingly, define

$$\tau(c) = \text{sum}(\tau_j : j \text{ in } I(c)). \quad (33)$$

Then, one can replace, in the limiting approximating sense of Theorem 1, $\text{meanconc}_4(\{\text{if } b_1, \text{ then } a_1, \ldots, \text{ if } b_m, \text{ then } a_m\}; \text{ if f, then e})(\underline{t})$ by

$$\text{meanconc}_2(\{\text{if } \beta, \text{ then } \alpha\}; \text{ if f, then e})(t) = Q(t) / R(t); \quad (34)$$

where

$$Q(t) = [\tau(\beta) \cdot \tau(e\&f\&\alpha)/\tau(\alpha)] \cdot t + [\tau(\beta) \cdot \tau(e\&f\&\alpha'\&\beta)/\tau(\alpha'\&\beta)] \cdot (1\text{-}t) + \tau(e\&f\&(\beta' \neg \omega_q)); \quad (35)$$

$$R(t) = [\tau(\beta) \cdot \tau(f\&\alpha)/\tau(\alpha)] \cdot t + [\tau(\beta) \cdot \tau(f\&\alpha'\&\beta)/\tau(\alpha'\&\beta)] \cdot (1\text{-}t) + \tau(f\&(\beta' \neg \omega_q)) \quad (36)$$

noting the relations from eqs.(32), (33),

$$\tau(e\&f\&(\beta' \neg \omega_q)) = \tau(e\&f) - \tau(e\&f\&\beta) - \tau_q, \quad \blacksquare$$

Corollary 1. (New application of Theorem 2 to modified transitivity problem in eq.(7))

Consider the modified transitivity problem given in eq.(7), where events a, b_j, b, c_j, c are all interpreted as before in eq.(3). (Here, we no longer require any use of eq.(4).) In general, there are $q = 2 \cdot (2^m\text{-}1) + 5$ relative atoms here and all assumptions of Theorem 1 hold, including eq.(24). Thus, making again the assumptions of Theorem 2, applied to the problem here, it follows that the approximating results, as well as the computations in eqs. (34)-(36) all hold here, where a common threshold t is determined from initially given s and t_j in eq.(7) (such as via a weighted average). In particular, letting $I_m = \{1, \ldots, m\}$,

$$\beta = b \vee c_1 \vee \ldots \vee c_m \, ;$$

$$\alpha = (a \,\&\, b \,\&\, c) \vee \left[\vee_{\varnothing \neq K \subseteq I_m} \left([\&_{j \text{ in } K}(b_j')] \,\&\, [\&_{i \text{ in } I_m \neg K}(b_i)]\right)\right];$$

$$\alpha' \& \beta = (a' \& b) \vee (b_1' \& c_1) \vee \ldots \vee (b_m' \& c_m) \, ;$$

$$\beta' = b' \& c_1' \& \ldots \& c_m' \, ; \quad (37)$$

where here e = a, f = c, so that

$$e \,\&\, f \,\&\, \alpha = (a \,\&\, c) \,\&\, \left[\vee_{\varnothing \subseteq K \subseteq I_m} \left([\&_{j \text{ in } K}(b_j')] \,\&\, [\&_{i \text{ in } I_m \neg K}(b_i)] \right)\right],$$

$$= a \& c \& (c_1 \Rightarrow b_1) \& \ldots \& (c_m \Rightarrow b_m);$$

$$e \& f \& \alpha' \& \beta = a \& b' \& c \, ; \; e \& f \& \beta' = \varnothing;$$

$$f \& \alpha = (a \& b \& c) \vee \left[\vee_{\varnothing \neq K \subseteq I_m} \left(c \,\&\, [\&_{j \text{ in } K}(b_j')] \,\&\, [\&_{i \text{ in } I_m \neg K}(b_i)]\right)\right];$$

$$f \& \alpha' \& \beta = a' \& b \& c \;\vee\; b' \& c \, ; \; f \& \beta' = \varnothing. \quad (38)$$

In turn, using the definition in eq.(33) in a straightforward way, eqs.(37) and (38) lead to the corresponding additive computations for $\tau(\alpha)$, $\tau(\alpha'\&\beta)$, $\tau(e\&f\&\alpha)$, $\tau(e\&f\&\alpha'\&\beta)$, $\tau(e\&f\&\beta')$, $\tau(f\&\alpha)$, $\tau(f\&\alpha'\&\beta)$, $\tau(f\&\beta')$, etc., yielding the full evaluation of eqs.(35) and (36), and hence eq.(34). ■

4. Choice of attributes and other issues in applying CRANOF

In implementing either the transitivity or the less restrictive modified transitivity approach, one must choose the most appropriate system features or attributes to be able to compare normal or attacked cyber states relative to these categories. One possible set is furnished in [15], where, in effect the relevant attributes under consideration for detection of intrusion were associated with *sendmail* system traces, with corresponding domains given in possible post-processing percentages of abnormal sequences of system traces. These include various types of *sscp* (*sunsendmailcp*), *syslog-remote*, *syslog-local*, and *decode*.

Additional issues to be addressed in future work include: (i) strategies for obtaining the empirical distributions necessary for obtaining initial s and the t_j in eq.(7), as discussed in Corollary 1; (ii) quantitative sensitivity analysis of estimated correlation levels to choices of sets of attributes and number of attributes; and (iii) complexity problems arising from the number of alternative states considered for pairwise comparison with current state.

Acknowledgements

The first author wishes to express his appreciation to the Chief of Naval Operations (CNO) Strategic Studies Group,

Newport, RI for their partial support of this work. The second two authors wish to express their appreciation for support of this work by the Office of Naval Research (ONR) In-House Laboratory Independent Research program (ILIR), FY01, Project No. ZA014.

References

[1] E.W. Adams, "On the logic of high probability", *Journal of Philosophical Logic*, vol. 15, 1986, pp. 255-279.

[2] E.W. Adams, "Four probability-preserving properties of inferences", *Journal of Philosoph. Logic*, vol. 25, 1996, pp. 1-24.

[3] D. Bamber, "Entailment with near surety of scaled assertions of high conditional probability", *Journal of Philosoph. Logic*, vol. 29, 2000, pp. 1-74.

[4] D. Bamber & I. R. Goodman, "New uses of second-order probability in estimating critical probabilities in Command & Control decision-making", *Proc. 2000 Command & Control Research & Technology Symposium* (Naval Postgraduate School, Monterey, CA. June 26-28, 2000); 53 pp., <http://www.dodccrp.org/2000CCRTS/cd/html/pdf_papers/Track_4/124.pdf>.

[5] D. Bamber, I.R. Goodman & H.T. Nguyen, "Extension of the concept of propositional deduction from classical logic to probability: an overview of probability-selection approaches" *Information Sciences*, vol. 131, 2001, pp. 195-250.

[6] D. Bamber, I.R. Goodman, W.C. Torrez & H.T. Nguyen, "Complexity reducing algorithm for near optimal fusion (CRANOF) with applications to tracking and information fusion", *Proc. Signal Processing, Sensor Fusion & Target Recognition X* (I. Kadar, ed.), SPIE vol. 4380 (AeroSense 2001, Orlando, FL, April 16-20, 2001), to appear 2001.

[7] G. de Cooman & P. Walley, *The Imprecise Probabilities Project*, <http://ippserv.rug.ac.be/home/ipp.html>.

[8] DARPA Information Assurance & Survivability, I: Strategic Intrusion Assessment, <http://www.afrlsn.afrl.af.mil/IA&S_topics.html>.

[9] M. Goldszmidt, P. Morris & J. Pearl, "A maximum entropy approach to nonmonotonic reasoning", *IEEE Trans. on Pattern Analysis & Machine Intelligence*, vol. 15(3), 1993, pp. 220-232.

[10] I.R. Goodman & D. Bamber, "An approach to extending classical entailment to a probabilistic setting: the fixed threshold case", to be submitted (2001).

[11] I.R. Goodman & H.T. Nguyen, Application of conditional and relational event algebra to the defining of fuzzy logic concepts", *Proc. Signal Processing, Sensor Fusion & Target Recognition VIII* (I. Kadar, ed.), SPIE vol. 3720 (AeroSense '99, Orlando, FL, April 5-7, 1999), pp. 25-36.

[12] I.R. Goodman & H.T. Nguyen, "Probability updating using second order probabilities and conditional event algebra", *Information Sciences*, vol. 121, 1999, pp. 295-347.

[13] J.B. Kadane (ed.), *Robustness of Bayesian Analysis*, North-Holland, Amsterdam, 1984.

[14] S. Kraus, D. Lehmann & M. Magidor, "Nonmonotonic reasoning, preferential models, and cumulative logics", *Artificial Intelligence*, vol. 44, 1990, pp. 167-207.

[15] W. Lee and S.J. Stolfo, *Data Mining Approaches for Intrusion Detection*, Report from Computer Science Dept., Columbia University, New York City, as of May, 2001 (22 pages); <http://www.cs.columbia.edu/~sal/hpapers/USENIX/usenix.html>.

[16] L.R. Pericchi, "Sets of prior probabilities and bayesian robustness" (7 pages), <http://ippserv.rug.ac.be/documentation/robust/robust.html>. (See also [7].)

[17] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, CA, 1988.

[18] Strategic Studies Group (SSG), Naval War College, Newport, RI, <http://www.nwc.navy.mil/ssg/>.

[19] W. Torrez and W. Yssel, "Associating microwave radar tracks with Relocatable Over-the-Horizon Radar (ROTHR) tracks using the Advanced Tactical Workstation", *Proceedings of the 32nd Asilomar Conference On Signals, Systems, and Computers, Vol. 1* (Pacific Grove, CA, Nov. 1-4, 1998), pp. 618-622.

[20] A. Wald, *Statistical Decision Functions*, Wiley, New York, 1950.

[21] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall, London, 1991.